

O'Brien N, McDonald S, Araújo-Soares V, Lara J, Errington L, Godfrey A, Mathers JC, Meyer TD, Rochester L, White M, Sniehotta FF. [The features of interventions associated with long-term effectiveness of physical activity interventions in adults aged 55 to 70 years: a systematic review and meta-analysis](#). *Health Psychology Review* 2015, 9(4), 417-433.

**Copyright:**

This is an Accepted Manuscript of an article published by Taylor & Francis in *Health Psychology Review* on 17/02/2015, available online: <http://dx.doi.org/10.1080/17437199.2015.1012177>

**DOI link to article:**

<http://dx.doi.org/10.1080/17437199.2015.1012177>

**Date deposited:**

29/10/2015

**Embargo release date:**

17 February 2016



This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](#)

**The features of interventions associated with long-term effectiveness of physical activity interventions in adults aged 55 to 70 years: a systematic review and meta-analysis**

Nicola O'Brien\*, Institute of Health & Society, Newcastle University, Baddiley-Clark Building, Richardson Road, Newcastle upon Tyne, NE2 4AX, UK. Tel: +44 (0)191 2088915; [nicki.o'brien@ncl.ac.uk](mailto:nicki.o'brien@ncl.ac.uk)

Suzanne McDonald, Institute of Health & Society, Newcastle University, Baddiley-Clark Building, Richardson Road, Newcastle upon Tyne, NE2 4AX, UK. Tel: +44 (0)191 2087361; [s.mcdonald1@ncl.ac.uk](mailto:s.mcdonald1@ncl.ac.uk)

Vera Araújo-Soares, Institute of Health & Society, Newcastle University, Baddiley-Clark Building, Richardson Road, Newcastle upon Tyne, NE2 4AX, UK. Tel: +44 (0)191 2086083; [vera.araujo-soares@ncl.ac.uk](mailto:vera.araujo-soares@ncl.ac.uk)

Jose Lara, Human Nutrition Research Centre, Institute for Ageing and Health, Newcastle University, Biomedical Research Building, Campus for Ageing and Vitality, Newcastle upon Tyne, NE4 5PL, UK. Tel: +44 (0)191 2081141; [jose.lara@ncl.ac.uk](mailto:jose.lara@ncl.ac.uk)

Linda Errington, Walton Library, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK. Tel: +44 (0)191 2088108; [linda.errington@ncl.ac.uk](mailto:linda.errington@ncl.ac.uk)

Alan Godfrey, Clinical Ageing Research Unit, Newcastle University, Campus for Ageing and Vitality, Newcastle upon Tyne, NE4 5PL, UK. Tel: +44 (0)191 2081245; [alan.godfrey@ncl.ac.uk](mailto:alan.godfrey@ncl.ac.uk)

Thomas D Meyer, Department of Psychiatry and Behavioral Sciences, University of Texas at Houston, 1941 East Road, BBSB 3118, Houston, Texas 77054. Tel: +1 (713) 4862563; [Thomas.D.Meyer@uth.tmc.edu](mailto:Thomas.D.Meyer@uth.tmc.edu)

Lynn Rochester, Clinical Ageing Research Unit, Newcastle University, Campus for Ageing and Vitality, Newcastle upon Tyne, NE4 5PL, UK. Tel: +44 (0)191 2081291; [lynn.rochester@ncl.ac.uk](mailto:lynn.rochester@ncl.ac.uk)

John C Mathers, Human Nutrition Research Centre, Institute for Ageing and Health, Newcastle University, Biomedical Research Building, Campus for Ageing and Vitality, Newcastle upon Tyne, NE4 5PL, UK. Tel: +44 (0)191 2081133; [john.mathers@ncl.ac.uk](mailto:john.mathers@ncl.ac.uk)

Martin White, Fuse, UKCRC Centre for Translational Research in Public Health, Institute of Health & Society, Newcastle University, Baddiley-Clark Building, Richardson Road, Newcastle upon Tyne, NE2 4AX, UK. Tel: +44 (0)191 2086275; [martin.white@ncl.ac.uk](mailto:martin.white@ncl.ac.uk)

Falko F Sniehotta, Fuse, UKCRC Centre for Translational Research in Public Health, Institute of Health & Society, Newcastle University, Baddiley-Clark Building, Richardson Road, Newcastle upon Tyne, NE2 4AX, UK. Tel: +44 (0)191 2083815; [falko.sniehotta@ncl.ac.uk](mailto:falko.sniehotta@ncl.ac.uk)

\*Corresponding author

This work is part of the LiveWell programme. LiveWell is supported by the Lifelong Health and Wellbeing initiative (LLHW), under Grant number G0900686. The LLHW initiative is a funding collaboration between the UK Research Councils and Health Departments. The funding partners are: Biotechnology and Biological Sciences Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Medical Research Council, Chief Scientist Office of the Scottish Government Health Directorates, National Institute for Health Research /The Department of Health, The Health and Social Care Research & Development of the Public Health Agency (Northern Ireland), and Wales Office of Research and Development for Health and Social Care, Welsh Assembly Government.

## Abstract

Content, delivery and effects of physical activity (PA) interventions are heterogeneous. There is a need to identify intervention features (content and delivery) related to long-term effectiveness. Behaviour change techniques (BCTs) and modes of intervention delivery were coded in 19 randomised controlled trials included in a systematic review of PA interventions for adults aged 55-70 years, published between 2000 and 2010, with PA outcomes  $\geq 12$  months after randomisation; protocol registration: PROSPERO CRD42011001459. Meta-analysis, moderator analyses and meta-regression were conducted. Meta-analysis revealed that interventions were effective in promoting PA compared with no/minimal intervention comparators ( $d=0.29$ , 95% CI=0.19 to 0.40,  $I^2=79.8\%$ , Q-value=89.16 (df=18,  $p<0.01$ )). Intervention features often concurred and *goal setting* was the most commonly used BCT. Subgroup analyses suggested that interventions using the BCT *feedback* may be more effective, whilst interventions using *printed materials* or the BCTs *information on where and when to perform the behaviour* and *information on consequences of behaviour to the individual* may be less effective. Meta-regression revealed that neither the number of BCTs nor self-regulatory BCTs significantly related to effect size. *Feedback* appears to be a potentially effective candidate BCT for future interventions promoting long-term PA. Considering concurrence of intervention features alongside moderator analyses is important.

**Keywords:** Behaviour change techniques, modes of delivery, physical activity intervention, meta-analysis, moderators

## Introduction

The health benefits of engaging in regular physical activity (PA) are numerous and well documented (Kohl et al., 2012). Moreover, there is a positive association between physical

inactivity and the prevalence of age-related disease and disability (Lee et al., 2012).

International guidelines recommend that each week adults, including those aged 65 years and over, should do at least 150 minutes of moderate-intensity aerobic PA, 75 minutes of high-intensity aerobic PA, or a combination thereof (World Health Organization, 2010). Despite this recommendation, less than 10% of middle to older age adults meet current recommendations when using objective measures of PA (Tucker, Welk, & Beyler, 2011).

Short to medium-term, moderate positive effects of PA interventions in middle to older age have been identified through systematic reviews (Conn, Hafdahl, & Mehr, 2011; Foster, Hillsdon, & Thorogood, 2005). Extending this evidence base, we conducted a systematic review with meta-analyses to quantify the longer-term effects ( $\geq 12$  months) according to the method of PA assessment and the duration of follow-up, and concluded that there is evidence for improvements in PA in adults aged 55 to 70 years at 12 months but not beyond (Hobbs et al 2013). However, like other reviews of behavioural interventions (Dombrowski et al., 2012; Michie, Abraham, Whittington, McAteer, & Gupta, 2009), our review identified moderate to high levels of heterogeneity in effect, which begs the question of what factors may be moderating the effects.

Behavioural interventions are complex and typically contain multiple, varying features and components, and it has been suggested that these features may help to explain the observed heterogeneity. To explore whether these features moderate the intervention effects identified by systematic reviews, one first needs to reliably identify the features which were used in each intervention. Development of taxonomies of behaviour change techniques (BCTs) (Abraham & Michie, 2008; Michie et al., 2011; Michie et al., 2013) has provided a framework to reliably code the content of behavioural interventions. In addition to synthesising the evidence for interventions that target a particular health behaviour in a particular population, health psychology research has also utilised BCT taxonomies to investigate the moderating effects of individual and empirically or theoretically derived clusters of intervention features (Bird et al.,

2013; Dombrowski et al., 2012; Dusseldorp, van Genugten, van Buuren, Verheijden, & van Empelen, 2013; Greaves et al., 2011; Michie et al., 2009; Taylor, Conner, & Lawton, 2012; Webb, Joseph, Yardley, & Michie, 2010). Consequently evidence is accumulating for the most and least effective intervention features, in turn, informing the development of future interventions (Michie & Johnston, 2012). However, a debate exists surrounding the methods used to synthesise the evidence of the impact of behavioural features, such as BCTs (Michie, Johnson, & Johnston, 2014; Peters, de Bruin, & Crutzen, 2013). In particular, ensuring that analyses consider and account for the concurrence of BCTs (i.e., the simultaneous use of more than one BCT in an intervention) has been highlighted due to the potentially synergistic effects of combinations of BCTs.

In addition to the potential moderating role of BCTs on intervention effectiveness, the mode of intervention delivery may also play an important role. Intervention modes of delivery, in terms of the provider, format, setting and intensity, can be coded in line with guidance on the minimal intervention characteristics that should be reported in public health interventions (Davidson et al., 2003).

This systematic review synthesises evidence for long-term effects of PA interventions in adults aged 55 to 70 years. The aims of the analyses are twofold: 1) to identify the concurrence and potential inter-relationships between intervention features (modes of delivery and BCTs); and 2) to explore associations between intervention features and intervention effectiveness.

## **Method**

### *Study selection criteria and search strategy*

This systematic review adhered to a registered protocol (Hobbs et al., 2011). Full details of the inclusion and exclusion criteria, and the search strategy are provided in an earlier paper reporting intervention effects according to PA outcomes and duration of follow-up assessment (Hobbs et al., 2013). Briefly we included RCTs of interventions assessing and reporting PA

behaviour  $\geq 12$  months after randomization. Interventions were compared with no intervention, minimal intervention or usual care control comparisons. Included trials studied free-living, healthy participants or those 'at risk' of chronic disease with a mean or median age of 55 to 70 years. Publications of any language with an English language abstract and from one of the 'most developed countries' within the United Nations index (United Nations Development Program, 2011) were considered for inclusion. Trials of participants recruited on the basis of taking a particular medication or with a pre-existing chronic or acute medical condition were excluded. Trials of laboratory-based exercise, performance training and those that reported physiological rather than behavioural measures of PA were also excluded. Twelve electronic databases were searched for articles published between January 2000 and November 2010 using search terms relating to PA, middle- to older-aged people and RCT (see supplementary file 1 for OVID Medline search strategy). Title, abstract and full text screening was conducted independently by two reviewers and, when necessary, in consultation with a third reviewer. Reference lists of included publications and reviews of PA interventions were searched for additional publications.

#### *Data extraction*

Outcome data from each included publication were extracted by one reviewer and independently checked by another (Kappa = 0.86); a third reviewer was consulted to resolve discrepancies. Authors were contacted for missing data and additional published and unpublished intervention materials, manuals and descriptions to facilitate the accuracy of coding intervention features.

The following intervention features were coded: BCTs employed and modes of delivery (intervention provider, format, setting and intensity (Davidson et al., 2003)). The CALO-RE taxonomy of BCTs (Michie et al., 2011) was used to code BCTs as it was specifically developed for use with PA and dietary interventions. The presence of each of the 40 BCTs was judged against each BCT definition provided in the taxonomy. When interventions addressed

lifestyle behaviours, such as diet or smoking, in addition to PA behaviour, then only BCTs that were specific to changing PA were coded as present. BCTs in the control conditions were also coded. To isolate the effect of an individual BCT as part of the intervention, if a BCT was identified in both the intervention and control conditions, then the BCT was only coded as being present in the intervention when it was used in a more intensive way than in the control condition. That is, the BCT was coded when it was used in a longer total period of active intervention contact and/or administered using more intensive, active modes of delivery (e.g. face-to-face rather than in printed material). Where the theoretical basis of the design or delivery of the intervention was stated, the theory/model(s) used were extracted. Intervention intensity was coded in terms of the length of the intervention period and the number of intervention contacts. The intervention period was defined as beginning when participants were first exposed to the intervention (i.e., when the first intervention contact occurred) and ending when the intervention finished (i.e., after the last intervention contact occurred). Intervention contacts included sessions to boost or promote maintenance of intervention effects but not those where only study outcomes were assessed. The intervention features of each included trial were coded independently by two of three reviewers, all of whom had been trained to code by one of the authors of the taxonomy and were experienced using the approach in other systematic reviews (NH, SMcD & VAS) (Kappa = 0.88); discrepancies were resolved in consultation with the other reviewer.

Trial quality was assessed using the Cochrane risk of bias tool (Higgins & Green, 2011). For each of the seven bias criteria (adequate sequence generation, allocation concealment, blinding (participants, personnel and assessors), incomplete outcome data addressed and free of selective outcome reporting) trials were categorized as low risk of bias scoring 0, or unclear or high risk of bias scoring 1. Therefore, an overall score from 0 to 7 was calculated for each trial, with higher scores indicating greater risk of bias.

### *Data Analysis*



The effect of interventions on PA behaviour when compared with the control condition was examined. To identify the potential moderating effects of intervention features, data reported at the follow-up time point which occurred immediately at the end of the intervention period were used for meta-analyses.

If trials assessed PA using multiple methods, then data collected objectively rather than by self-report, on total PA rather than specific PA domains and reported as duration rather than metabolic equivalents were used for analyses. When a trial reported PA both on a continuous and a dichotomous measure, the continuous measure was used for analysis to maximise statistical power and retain more information about the construct. Data from intention-to-treat analyses were used when reported.

When trials reported change scores from baseline, final values were computed where possible or requested from authors. When only the median and range values were reported, mean and standard deviation values were requested from authors. When mean and standard deviation values were unavailable, missing data were imputed using the median instead of the mean and by estimating the standard deviation from the standard error, inter-quartile range or range (Higgins & Green, 2011; Hozo, Djulbegovic, & Hozo, 2005).

The following rules were applied when trials had multiple intervention arms: the PA intervention arm was compared with the control condition when the other intervention arm targeted diet; and the most intensive PA intervention arm was compared with the control condition when the other intervention arms also targeted PA. The most intensive intervention was defined as the intervention which included the most intervention contacts, modes of intervention delivery or BCTs.

Standardised mean differences (SMDs) or odds ratios with 95% confidence intervals (CIs) were calculated for all trials. Results were pooled using a random effects model (inverse-variance approach based on weighted SMDs and odds ratios) using RevMan software (version 5.1) (Review Manager, 2011). Random effects models were used to estimate the mean of a

distribution of effects across the included trials of behavioural interventions. By their very nature, such interventions are heterogeneous in terms of modes of delivery and content (BCTs). Trials with dichotomous outcomes were merged with trials with continuous outcomes using Comprehensive Meta-Analysis software (CMA) (version 2.2) (Borenstein, Hedges, Higgins, & Rothstein, 2005) to produce SMDs for each trial, which are equivalent to Cohen's  $d$  (Cohen, 1988).

The Cochran Q test was used to investigate the presence or absence of statistical heterogeneity (i.e., true differences) in effect sizes across trials. The  $I^2$  test statistic was used to indicate the percentage of total variation explained by any identified heterogeneity (Higgins & Green, 2011). Sensitivity analyses were conducted to test the robustness of the findings when trials with missing data, high attrition and/or high risk of bias were removed from the analysis. As previously reported (Hobbs et al., 2013), there was no statistical evidence that publication bias affected the results.

Exploratory analyses of the concurrence of intervention features were conducted using incidence and concurrence matrices developed using Minitab (version 17). The incidence matrix showed whether each intervention feature (modes of delivery and BCTs) was used or not used in each included trial. The concurrence matrix, which is the transpose of the incidence matrix multiplied by the incidence matrix<sup>1</sup>, showed how often each feature was used together with any other feature, i.e., the concurrence or co-occurrence of pairs of features. Individual intervention features were included in these exploratory analyses only when they had been used in at least 25% of included trials but not more than 75%. These cut-offs were used to ensure there was sufficient variability in the number of trials where each intervention feature

---

<sup>1</sup> An incidence matrix shows the relationship between two factors, which in this case is each trial and each intervention feature. The incidence matrix uses '1' to show when a feature was used in a trial and '0' to show when a feature was not used. This matrix is multiplied by its transpose (i.e., where the incidence matrix rows are interchanged with columns) which produces the concurrence matrix.

was coded as present or absent for the analyses. The percentage of trials employing all possible combinations of pairs of intervention features was calculated.

Random effects subgroup analyses with Q statistic tests of subgroup differences were conducted using CMA software. These univariate moderator analyses were applied to investigate the association between use of specific intervention features and intervention effectiveness. Trials were further categorised as lower or higher risk of bias using a median cut-off split of the overall risk of bias score (median = 5). Random effects univariate meta-regression analysis was conducted to explore whether the number of BCTs and the number of self-regulatory BCTs employed in each intervention was associated with intervention effectiveness.

## **Results**

Nineteen trials were included in the meta-analysis. The number of publications included and excluded at each stage, and reasons for exclusion are shown in supplementary file 2.

Supplementary file 3 displays the characteristics of included trials, which includes information on the study population; PA outcome data and assessment time; attrition rates; intervention materials coded in addition to the text in the publication; modes of delivery and BCTs used in the intervention and control conditions; and risk of bias score. In total, trials reported on 10,423 participants with a mean age of 60.7 years (SD = 6; range 55.0 to 67.6) and 64% of whom were female. Trials were conducted in the USA, Europe, New Zealand, Japan, Australia and Canada.

### *Modes of Delivery*

A detailed description of the intervention modes of delivery is reported in our earlier publication (Hobbs et al., 2013). In summary, the majority of intervention providers were health care professionals (16 of 19), one intervention was delivered by the participant under instruction, and the intervention provider was unclear in the remaining two trials. The majority of interventions (16 of 19) were, at least in part, delivered on a face-to-face basis on an

individual and/or group basis. Eleven of the trials with a face-to-face component were also delivered using printed materials and/or the telephone. The three trials with no face-to-face component were delivered via printed material, the internet and/or the telephone. Fourteen trials were delivered in a home setting, at least in part, because they were delivered wholly or in part over the telephone or internet. On average, the intervention period, including the principal intervention and any maintenance intervention period, was 11 months (SD = 7.2; range 1 to 24) with 29 contacts (SD = 42; range 1 to 164).

### *Behaviour Change Techniques*

A mean of 9 BCTs per intervention was coded (SD = 6; range 1 to 29). The most common BCTs used were *behavioural goal setting* (n=13), *self-monitoring of behaviour* (n=13), *graded tasks* (n=11), *feedback* (n=11), *barrier identification* (n=10) and *information on consequences of behaviour to the individual* (n=10). Interventions for which additional intervention material was provided by authors (n=7) had a mean of 13 BCTs coded compared with 7 BCTs in interventions for which additional materials were not provided.

### *Theoretical Basis of the Intervention*

Ten trials reported employing one or more behaviour change theory or model to design and/or deliver the intervention. The most common theories or models reported were the Transtheoretical Model (n=7) and Social Cognitive Theory (n=4).

### *Overall Trial Effectiveness*

The included trials reported PA outcomes as dichotomous measures (n= 3) and continuous measures (n=16). Data reported at 12 month post-randomisation follow-up time point were used for 14 trials, data at 18 month follow-up time point for two trials, data at 24 month follow-up time point for two trials, and data at 36 month follow-up for one trial. Comparing intervention and control conditions, a small positive effect was identified in favour of the interventions (Figure 1: SMD=0.29, 95% CI=0.19 to 0.40,  $k=19$ ). Statistical heterogeneity was high with an  $I^2$  of 79.8% and a Q-value of 89.16 (df=18,  $p<0.01$ ).

Figure 1 about here

Sensitivity analysis removing trials for which missing mean and standard deviation values were imputed or estimated reduced the size of the intervention effect (SMD=0.21, 95% CI=0.13 to 0.28,  $k=15$ ) and heterogeneity ( $I^2 = 54.9\%$ ,  $Q=31.07$  ( $df=14$ ,  $p<0.01$ )). However, sensitivity analysis removing additional trials with attrition rates over 30% and/or high risk of bias did not alter the effect size (SMD=0.22, 95% CI=0.13 to 0.31,  $k=12$ ) or heterogeneity ( $I^2 = 59.1\%$ ,  $Q=26.92$  ( $df=11$ ,  $p<0.01$ )) any further (see supplementary file 3 for data on trials with missing data, high attrition and/or high risk of bias).

#### *Exploratory Concurrence and Moderator Analyses of Intervention Features*

Each feature included in the exploratory analyses had been coded as present in at least 25% of trials (5 of 19) but not more than 75% (15 of 19). Thus, the following BCTs were not included: *shaping, anticipated regret, fear arousal, self-talk, imagery, stress management, time management, communication skills and anticipation of future rewards* ( $n=0$ ); *information about others' approval, reward effort, prompt generalisation, behavioural contract and identification as role model* ( $n=1$ ); *focus on past success and environmental restructuring* ( $n=2$ ); *information on consequences of behaviour in general and reward success* ( $n=3$ ); *normative information, action planning, review outcome goals, prompts/cues, social comparison, relapse prevention and motivational interviewing* ( $n=4$ ). The following modes of delivery were also excluded from analysis due to a lack of variability in occurrence: health professional as intervention provider ( $n=16$ ) and internet as intervention format ( $n=2$ ).

#### *Concurrence of Features*

Using the included intervention features, incidence and concurrence matrices were developed to explore how often each intervention feature was coded as being used and how often it occurred in combination with another (Table 1). The percentage of trials employing all possible combinations of pairs of intervention features was calculated as an indication of concurrence.

Table 1 about here

One hundred percent of the time, all interventions that used the BCT *outcome goal setting* were delivered at home, whilst all interventions that used the BCT *barrier identification* also used the BCT *self-monitoring of behaviour*. When the BCT *behavioural goal review* was used, the interventions were always delivered at home and used the BCTs *behavioural goal setting* and *follow-up prompts*. When the BCT *model/demonstrate the behaviour* was used, the interventions always used the BCTs *graded tasks* and *instruction on how to perform the behaviour*. When the BCT *prompt practice* was used, interventions also always used the BCT *graded tasks*. Finally, when the BCT *follow-up prompts* was used, interventions also always used the BCT *behavioural goal setting*. All interventions that were delivered on the telephone were also coded as being delivered at home, at least in part.

Table 2 about here

#### Moderating Effects of BCTs used

Table 2 displays which intervention features were used in each of the included trials. Trials are ordered from most to least effective using the trial SMD. The trial sample size and risk of bias (high or low) are also presented. Modes of delivery and BCTs are grouped separately.

Subgroup analyses comparing the effect size estimate for interventions using each feature compared with interventions not using the feature are presented. Features are ordered within their respective group of modes of delivery or BCTs by the size of the difference in effect size estimate between interventions using and not using each feature. Use of the BCT *feedback* was the only feature found to be associated with being more effective; interventions using *feedback* were significantly more effective than interventions that did not use *feedback* ( $Q=4.14$ ,  $p=0.042$ ). In contrast, interventions that used *printed materials* as mode of delivery, and the BCTs *information on where and when to perform the behaviour*, and *information on consequences of behaviour to the individual* were found to be associated with being less

effective (*printed materials*:  $Q=9.47$ ,  $p=0.002$ ; *information on where and when to perform the behaviour*:  $Q=6.06$ ,  $p=0.014$ ; and *information on consequences of behaviour to the individual*:  $Q=10.52$ ,  $p=0.001$ ).

#### *Moderating Effects of the Number of BCTs used*

Univariate meta-regression analysis found no significant association between the number of BCTs used and intervention effectiveness (beta = -0.003, 95%CI: -0.009 to -0.002;  $Q = 1.32$  ( $df=1$ ,  $p=0.25$ )). BCTs intrinsic to self-regulation theory (*goal setting, goal review, self-monitoring, feedback*) were some of the most commonly coded BCTs in interventions; therefore, the association between the number of self-regulatory BCTs used and intervention effectiveness was also explored; univariate meta-regression analysis found no significant association (beta = 0.012, 95%CI: -0.025 to 0.049;  $Q = 0.42$  ( $df=1$ ,  $p>0.05$ )).

## **Discussion**

This systematic review and meta-analysis is the first to explore the concurrence and moderating effects of intervention features within 19 RCTs promoting long-term ( $\geq 12$  months) PA behaviour in adults aged 55 to 70 years. Using standardised and reliable methods to code the intervention features of BCTs (Michie et al., 2011) and modes of delivery (Davidson et al., 2003), we identified feature concurrence and found that many features were not used in isolation but occurred in combination. Particular features were associated with intervention effectiveness but the total number of BCTs or the number of self-regulatory BCTs did not moderate effects.

#### *Findings in Context and Implications*

There was no evidence that intervention setting is important for effectiveness; neither a healthcare or home setting were associated with effectiveness, which is in line with other findings from PA interventions (Greaves et al., 2011). The lack of a dose response relationship between the total number of BCTs used in interventions and intervention effectiveness is also

not a novel finding. The meta-analyses by Dombrowski et al (2012) and Taylor et al (2012) similarly concluded that using more BCTs does not necessarily lead to improved effectiveness. The potentially moderating effect of the number of self-regulatory BCTs was also tested but again no association was identified. Evidence for the role of self-regulatory techniques is mixed. Particular clusters of BCTs taken from self-regulation theory have been shown to be associated with PA intervention effectiveness (Greaves et al., 2011; Michie et al., 2011). Conversely, self-regulatory BCTs have been shown to be associated with lower levels of PA in interventions in older adults (French, Olander, Chisholm, & Mc Sharry, 2014). Our univariate moderator analyses found that use of the self-regulatory BCT *feedback* was associated with more effective interventions, which has been reported previously (Dombrowski et al., 2012; Greaves et al., 2011; Michie et al., 2009). The BCT *feedback* is defined in the CALO-RE taxonomy as providing the person with data about their own recorded behaviour (following the BCT *self-monitoring of behaviour*) or commenting on discrepancies between a person's behavioural performance and goals (linking with the BCT *behavioural goal setting*). Our concurrence analyses showed that *feedback* did not occur in combination with either of these BCTs 100% of the time; however, all 11 interventions which used *feedback* also used at least one of these other self-regulatory techniques. If the BCT *feedback* is related to more effective interventions but always concurs with at least one other self-regulatory technique, then the possible combined effects of self-regulatory techniques need to be further investigated.

Debate about methods used to synthesise evidence of the impact of behavioural features, such as BCTs (Michie et al., 2014; Peters et al., 2013), highlights the importance of analyses which consider, and account for, the concurrence of BCTs. If BCTs are coded as being used together in a given intervention, then conclusions based on univariate meta-regression analyses of the association between discrete BCTs and intervention effectiveness (Dombrowski et al., 2012; Taylor et al., 2012) may be confounded by the fact that the BCT was not used in isolation. For example, the moderator analyses in the present review revealed that interventions using *printed*



*materials* were less effective. However, the concurrence analyses showed that seven of the eight interventions that used *printed materials* also used the BCT *information on consequences of behaviour to the individual*, which was also shown to be less effective. Thus, it is possible that the presence or magnitude of the identified negative moderating effect of using *printed materials* is not true but rather a confound due to the BCT *information on consequences of behaviour to the individual* being used at the same time in 88% of possible cases. In fact, the BCT *information on consequences of behaviour* has previously been identified as a negative moderator of PA behaviour change; less effective PA interventions were associated with use of the technique (Dombrowski et al., 2012).

The other BCT shown to negatively moderate intervention effectiveness was the BCT *information on where and when to perform the behaviour*. This finding replicates that found by French et al (2014) in their review of BCTs associated with PA in older adults. This finding is counter-intuitive since one would anticipate that knowing about possible opportunities to be more physically active would be helpful in doing so. It is noted however, that the exploratory, post hoc nature of these moderator analyses means that spurious or chance findings are a possibility and therefore, these findings should be interpreted with caution. Nonetheless, the findings can be used to generate hypotheses for future research about which BCTs may be more or less effective. Such hypotheses can be used to inform the design of interventions where the effectiveness of particular intervention features can be investigated. For example, factorial trials would be a useful design through which intervention features could be investigated systematically to help isolate the effects of individual and combinations of BCTs. Moreover, with advances in statistical approaches to understand the combined effects of BCTs in interventions (Dusseldorp et al., 2013), future analyses based on evidence with sufficient statistical power have the potential to allow more sensitive analyses of BCT synergistic effects. This review focused on the long-term effectiveness of physical activity interventions and only included trials with at least a 12-month follow-up. This focus may help to explain some of the

discrepancies between the BCTs identified as being associated with effectiveness in this review and those identified in other reviews of the effectiveness of BCTs on physical activity, which considered shorter term outcomes (e.g., Bélanger-Gravel, Godin, & Amireault, 2011). It is worth highlighting, therefore, that together with evidence on the determinants of physical activity maintenance (Amireault, Godin, & Vézina-Im, 2012), the findings in this review are particularly informative for the design of interventions with potential long-term effectiveness.

### *Strengths and Limitations*

The coding of intervention features was conducted using reliable methods and a standardised and appropriate BCT taxonomy, the 40-item CALO-RE taxonomy which was specifically developed for use with PA and dietary interventions (Michie et al., 2011). BCTs were coded by health psychologists trained and experienced in coding intervention features, yielding high inter-rater reliability ( $Kappa = 0.88$ ). Where possible, intervention features were coded using all available intervention materials provided by authors. Noticeably, approximately twice as many BCTs were coded in trials for which additional intervention materials were provided compared with trials for which only the original publication was available. This finding mirrors that reported by Lorencatto et al (2012). The rigorous approach we adopted to code BCTs using whatever material was provided by authors meant that in some cases BCTs were coded from lengthy intervention manuals or doctoral theses, whereas in other cases BCTs were coded from short, concise intervention descriptions in a journal article. As a consequence, more BCTs may have been coded in trials for which additional materials were available (because there was significantly more text from which judgements could be made), and this may have introduced a systematic bias in subsequent analyses. However, it may also be the case that those trials for which additional material was available did use more BCTs because these interventions had been developed with a stronger basis in behaviour change science. With the available evidence, we cannot distinguish between these alternatives. Developments in open access publishing including widespread availability of supplementary online files, the advent of journal policies

that require authors to provide intervention manuals with the article, and guidance for the reporting of interventions (Hoffmann et al., 2014) are likely to diminish the adverse effects of article word limits. Future systematic reviews of behaviour change interventions will then be able to undertake more robust investigations of differences in the number BCTs coded between trials.

Another strength of this review was the careful consideration for the potential active content of no/minimal intervention, or usual care control conditions in the analytic strategy, which has been highlighted as an important limitation in previous analyses (Peters et al., 2013). A BCT was only coded as present in an intervention when it had not been identified in a control condition or when it had been used more intensively compared with the control condition. This approach was adopted to isolate the effect of an individual BCT as a unique feature of the intervention.

Sensitivity analyses confirmed the robustness of our meta-analysis findings. Even though the size of the intervention effect was reduced by the removal of trials for which missing data had been estimated or those with a higher risk of bias (i.e., SMD of 0.29 reduced to 0.21 and 0.22 respectively), the significance of the positive effect of interventions was retained. Nonetheless, the associations between intervention features and effectiveness can be further interpreted in light of trial risk of bias. As presented in Table 2, the trial by Babazono (2007) is the most effective trial with a low risk of bias. The intervention in this trial used the BCT *feedback*, which was shown to be associated with more effective interventions in the moderator analyses, but did not use *printed materials* or the BCTs *information on where and when to perform the behaviour* and *information on consequences of behaviour to the individual*, which were shown to be associated with less effective interventions. However, the trial by Hertogh (2010), which was the most effective trial in the review, similarly used *feedback* but not *printed materials* or the BCTs *information on where and when to perform the behaviour* and *information on consequences of behaviour to the individual*; yet this trial was rated as high risk of bias as the

methods of randomisation, allocation concealment and blinding could not be ascertained from the publication. It may be the case that the robustness of the identified moderating role of particular intervention features is unaffected by trial bias in the same way as was seen in the sensitivity analysis of the main meta-analyses. Nevertheless, both of these trials have limited sample sizes and therefore the overall findings would need to be replicated in larger trials.

A limitation of this review is that the findings are only applicable to interventions conducted with participants in countries categorised as being one of the ‘most developed countries’ within the United Nations index (United Nations Development Program, 2011); therefore, their generalisability to lower income countries needs to be ascertained.

A possible limitation of the meta-analysis is that in order to examine the potential moderating role of intervention features on effectiveness, all included trials were aggregated regardless of the way in which the PA behavioural outcome was measured (objectively or self-reported).

Inaccurate recall and social desirability may result in self-reported measures of PA being over- or under-estimated (Kowalski, Rhodes, Naylor, Tuokko, & MacDonald, 2012) which means that the intervention effectiveness of a trial reporting objective measures of PA using accelerometry data may not be comparable with that of a trial using self-reported data. In our earlier paper from this systematic review (Hobbs et al., 2013), we conducted meta-analyses according to the method of PA assessment and the duration of follow-up and concluded that when compared with controls at 12 months, interventions had a larger positive effect on pedometer measured step-count ( $SMD = 1.09$ ,  $n=4$ ) than on self-reported continuous outcomes ( $SMD = 0.19$ ,  $n=11$ ) or self-reported dichotomous outcomes ( $OR = 1.63$ ,  $n=3$ ). This suggests that self-reported measures may underestimate intervention effects; however this explanation warrants further enquiry.

Despite our attempt to explore concurrence of BCTs and modes of delivery, our analyses were limited by the relatively small number of trials. Multivariate analyses of combinations and clusters of BCTs could not be conducted due to a lack of statistical power; thus, concurrence of

features in pairs only was conducted. Another limitation that should be noted is that the reliability of our findings is dependent on intervention fidelity in terms of whether the intervention features coded are a true reflection of what actually occurred in the intervention. Fidelity checks would allow knowledge to be gathered about whether the intervention was delivered as planned, the extent to which the features were delivered, and the quality of delivery thereof. Only two trials reported on the independent assessment of intervention fidelity through audio and video recordings (Harting et al., 2006; King et al., 2007) and therefore this issue cannot be examined further.

Replication of these findings is needed when a larger set of trials is available. The intricacies of the moderating role of individual and clusters of particular intervention features on long-term effectiveness should be further scrutinised alongside data on feature concurrence. Moreover, the moderating roles of theory in intervention design and multiple behaviour focus (i.e., interventions targeting other behaviours in addition to PA) are additional hypotheses that could be explored in a larger data set.

### *Conclusion*

The long-term effectiveness of PA interventions in middle to older age adults was not associated with the total number of BCTs or the number of self-regulatory BCTs used. Moderator analyses suggested that interventions aiming to promote PA in this population should consider using the BCT *feedback* in order to enhance effects. However, many pairs of intervention features were used in concurrence making it difficult to isolate the effects of specific BCTs unambiguously.

### **Acknowledgements**

The authors thank all authors who provided further data and additional intervention materials.

## References

- Abraham, C., & Michie, S. (2008). A taxonomy of behaviour change techniques used in interventions. *Health Psychology, 27*, 379 - 387
- Amireault, S., Godin, G., & Vézina-Im, L.-A. (2012). Determinants of physical activity maintenance: a systematic review and meta-analyses. *Health Psychology Review, 7*(1), 55-91. doi: 10.1080/17437199.2012.701060
- Bélanger-Gravel, A., Godin, G., & Amireault, S. (2011). A meta-analytic review of the effect of implementation intentions on physical activity. *Health Psychology Review, 7*(1), 23-54. doi: 10.1080/17437199.2011.560095
- Bird, E. L., Baker, G., Mutrie, N., Ogilvie, D., Sahlqvist, S., & Powell, J. (2013). Behavior change techniques used to promote walking and cycling: a systematic review. *Health Psychology, 32*(8), 829-838. doi: 10.1037/a0032078
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). Comprehensive Meta-analysis Version 2. Englewood NJ Biostat.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erlbaum.
- Conn, V. S., Hafdahl, A. R., & Mehr, D. R. (2011). Interventions to increase physical activity among healthy adults: meta-analysis of outcomes. *American Journal of Public Health, 101*(4), 751-758. doi: 10.2105/AJPH.2010.194381
- Davidson, K. W., Goldstein, M., Kaplan, R. M., Kaufmann, P. G., Knatterud, G. L., Orleans, C. T., . . . Whitlock, E. P. (2003). Evidence-based behavioral medicine: what is it and how do we achieve it? *Annals of Behavioral Medicine, 26*(3), 161-171
- Dombrowski, S. U., Sniehotta, F. F., Avenell, A., Johnston, M., MacLennan, G., & Araujo-Soares, V. (2012). Effects of specific behaviour change techniques on behaviour and weight change: A systematic review. *Health Psychology Review, 6*(1), 7-32

- Dusseldorp, E., van Genugten, L., van Buuren, S., Verheijden, M. W., & van Empelen, P. (2013). Combinations of Techniques That Effectively Change Health Behavior: Evidence From Meta-CART Analysis. *Health Psychology*. doi: 10.1037/hea0000018
- Foster, C., Hillsdon, M., & Thorogood, M. (2005). Interventions for promoting physical activity. *Cochrane Database of Systematic Reviews*, CD003180
- French, D. P., Olander, E. K., Chisholm, A., & Mc Sharry, J. (2014). Which Behaviour Change Techniques Are Most Effective at Increasing Older Adults' Self-Efficacy and Physical Activity Behaviour? A Systematic Review. *Annals of Behavioral Medicine*. doi: 10.1007/s12160-014-9593-z
- Greaves, C., Sheppard, K., Abraham, C., Hardeman, W., Roden, M., Evans, P., . . . The, I. S. G. (2011). Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. *BMC Public Health*, 11(1), 119
- Harting, J., van Assema, P., van Limpt, P., Gorgels, T., van Ree, J., Ruland, E., . . . de Vries, N. K. (2006). Cardiovascular prevention in the Hartslag Limburg project: effects of a high-risk approach on behavioral risk factors in a general practice population. *Preventive Medicine*, 43(5), 372-378
- Higgins, J. P. T., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*: The Cochrane Collaboration.
- Hobbs, N., Godfrey, A., Lara, J., Errington, L., Meyer, T., Rochester, L., . . . Sniehotta, F. (2013). Are behavioral interventions effective in increasing physical activity at 12 to 36 months in adults aged 55 to 70 years? a systematic review and meta-analysis. *BMC Medicine*, 11(1), 75
- Hobbs, N., Godfrey, A., Sniehotta, F., Rochester, L., White, M., Errington, L., & Mathers, J. (2011). A systematic review of the maintenance effects of physical activity randomised

- controlled trials in healthy and ‘at risk’ older adults: LiveWell. *PROSPERO: International Prospective Register of Systematic Reviews*, CRD42011001459
- Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., . . . Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *British Medical Journal*, 348(7), 348:g1687. doi: 10.1136/bmj.g1687
- Hozo, S., Djulbegovic, B., & Hozo, I. (2005). Estimating the mean and variance from the median, range, and the size of a sample. *BMC Medical Research Methodology*, 5(1), 13
- King, A. C., Friedman, R., Marcus, B., Castro, C., Napolitano, M., Ahn, D., & Baker, L. (2007). Ongoing physical activity advice by humans versus computers: the Community Health Advice by Telephone (CHAT) trial. *Health Psychology*, 26(6), 718-727
- Kohl, H. W., Craig, C. L., Lambert, E. V., Inoue, S., Alkandari, J. R., Leetongin, G., & Kahlmeier, S. (2012). The pandemic of physical inactivity: global action for public health. *Lancet*, 380(9838), 294-305
- Kowalski, K., Rhodes, R., Naylor, P.-J., Tuokko, H., & MacDonald, S. (2012). Direct and indirect measurement of physical activity in older adults: a systematic review of the literature. *International Journal of Behavioral Nutrition and Physical Activity*, 9(1), 148
- Lee, I. M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., & Katzmarzyk, P. T. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet*, 380(9838), 219-229
- Lorencatto, F., West, R., Stavri, Z., & Michie, S. (2012). How Well Is Intervention Content Described in Published Reports of Smoking Cessation Interventions? *Nicotine & Tobacco Research*. doi: 10.1093/ntr/nts266



- Michie, S., Abraham, C., Whittington, C., McAteer, J., & Gupta, S. (2009). Effective techniques in healthy eating and physical activity interventions: A meta-regression. *Health Psychology, 28*, 690 - 701
- Michie, S., Ashford, S., Sniehotta, F. F., Dombrowski, S. U., Bishop, A., & French, D. P. (2011). A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. *Psychology & Health, 26*(11), 1479-1498. doi: 10.1080/08870446.2010.540664
- Michie, S., Johnson, B. T., & Johnston, M. (2014). Advancing cumulative evidence on behaviour change techniques and interventions: a comment on Peters, de Bruin, and Crutzen. *Health Psychology Review, 1-5*. doi: 10.1080/17437199.2014.912538
- Michie, S., & Johnston, M. (2012). Theories and techniques of behaviour change: Developing a cumulative science of behaviour change. *Health Psychology Review, 6*(1), 1-6. doi: 10.1080/17437199.2012.654964
- Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., . . . Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine, 46*(1), 81-95. doi: 10.1007/s12160-013-9486-6
- Peters, G.-J. Y., de Bruin, M., & Crutzen, R. (2013). Everything should be as simple as possible, but no simpler: towards a protocol for accumulating evidence regarding the active content of health behaviour change interventions. *Health Psychology Review, 1-14*. doi: 10.1080/17437199.2013.848409
- Review Manager. (2011). RevMan (Version 5.1). Copenhagen, Denmark: The Nordic Cochrane Centre, The Cochrane Collaboration.

- Taylor, N., Conner, M., & Lawton, R. (2012). The impact of theory on the effectiveness of worksite physical activity interventions: a meta-analysis and meta-regression. *Health Psychology Review*, 6(1), 33-73. doi: 10.1080/17437199.2010.533441
- Tucker, J. M., Welk, G. J., & Beyler, N. K. (2011). Physical activity in U.S.: adults compliance with the Physical Activity Guidelines for Americans. *American Journal of Preventive Medicine*, 40(4), 454-461. doi: 10.1016/j.amepre.2010.12.016
- United Nations Development Program. (2011). *Human Development Report 2011*. New York.
- Webb, L. T., Joseph, J., Yardley, L., & Michie, S. (2010). Using the Internet to Promote Health Behavior Change: A Systematic Review and Meta-analysis of the Impact of Theoretical Basis, Use of Behavior Change Techniques, and Mode of Delivery on Efficacy. *Journal of Medical Internet Research*, 12(1), e4. doi: 10.2196/jmir.1376
- World Health Organization. (2010). *Global recommendations on physical activity for health*. Geneva.

Table 1. Concurrence of pairs of intervention features coded as present in at least 25% but not more than 75% of trials

|                             |                                     | Total | A        | B       | C       | D                       | E       | F                       | G       | H        | I                      | J       | K                       | L       | M        | N       | O                      | P       | Q       | R                      | S       |
|-----------------------------|-------------------------------------|-------|----------|---------|---------|-------------------------|---------|-------------------------|---------|----------|------------------------|---------|-------------------------|---------|----------|---------|------------------------|---------|---------|------------------------|---------|
| Modes of Delivery           | A Telephone                         | 10    |          | 40<br>4 | 50<br>5 | <b>100</b><br><b>10</b> | 60<br>6 | 70<br>7                 | 40<br>4 | 60<br>6  | 60<br>6                | 40<br>4 | 70<br>7                 | 10<br>1 | 60<br>6  | 30<br>3 | 40<br>4                | 30<br>3 | 30<br>3 | 70<br>7                | 50<br>5 |
|                             | B Print material                    | 8     | 50<br>4  |         | 50<br>4 | 75<br>6                 | 88<br>7 | 63<br>5                 | 38<br>3 | 50<br>4  | 50<br>4                | 25<br>2 | 63<br>5                 | 25<br>2 | 75<br>6  | 50<br>4 | 38<br>3                | 25<br>2 | 13<br>1 | 63<br>5                | 38<br>3 |
|                             | C Healthcare                        | 10    | 50<br>5  | 40<br>4 |         | 70<br>7                 | 50<br>5 | 60<br>6                 | 30<br>3 | 40<br>4  | 60<br>6                | 30<br>3 | 60<br>6                 | 30<br>3 | 60<br>6  | 20<br>2 | 30<br>3                | 30<br>3 | 20<br>2 | 50<br>5                | 40<br>4 |
|                             | D Home                              | 14    | 71<br>10 | 43<br>6 | 50<br>7 |                         | 57<br>8 | 64<br>9                 | 36<br>5 | 57<br>8  | 57<br>8                | 36<br>5 | 65<br>11                | 21<br>3 | 71<br>10 | 29<br>4 | 43<br>6                | 29<br>4 | 29<br>4 | 64<br>9                | 57<br>8 |
| Behaviour Change Techniques | E Info consequence (individual)     | 10    | 60<br>6  | 70<br>7 | 50<br>5 | 80<br>8                 |         | 60<br>6                 | 40<br>4 | 50<br>5  | 50<br>5                | 30<br>3 | 70<br>7                 | 20<br>2 | 50<br>5  | 40<br>4 | 30<br>3                | 20<br>2 | 10<br>1 | 60<br>6                | 50<br>5 |
|                             | F Goal setting (behaviour)          | 13    | 54<br>7  | 38<br>5 | 46<br>6 | 69<br>9                 | 46<br>6 |                         | 31<br>4 | 62<br>8  | 62<br>8                | 38<br>5 | 69<br>9                 | 23<br>3 | 54<br>7  | 31<br>4 | 31<br>4                | 23<br>3 | 23<br>3 | 85<br>11               | 46<br>6 |
|                             | G Goal setting (outcome)            | 5     | 80<br>4  | 60<br>3 | 60<br>3 | <b>100</b><br><b>5</b>  | 80<br>4 | 80<br>4                 |         | 80<br>4  | 80<br>4                | 80<br>4 | 80<br>4                 | 0<br>0  | 80<br>4  | 20<br>1 | 40<br>2                | 40<br>2 | 20<br>4 | 80<br>4                | 60<br>3 |
|                             | H Barrier identification            | 10    | 60<br>6  | 40<br>4 | 40<br>4 | 80<br>8                 | 60<br>6 | 70<br>7                 | 40<br>4 |          | 70<br>7                | 40<br>4 | <b>100</b><br><b>10</b> | 10<br>1 | 60<br>6  | 30<br>3 | 30<br>3                | 10<br>1 | 20<br>2 | 70<br>7                | 60<br>6 |
|                             | I Graded tasks                      | 11    | 55<br>6  | 36<br>4 | 55<br>6 | 73<br>8                 | 45<br>5 | 73<br>8                 | 36<br>4 | 64<br>7  |                        | 36<br>4 | 82<br>9                 | 36<br>4 | 73<br>8  | 27<br>3 | 45<br>5                | 45<br>5 | 45<br>5 | 55<br>6                | 45<br>5 |
|                             | J Behavioural goal review           | 5     | 80<br>4  | 40<br>2 | 60<br>3 | <b>100</b><br><b>5</b>  | 60<br>3 | <b>100</b><br><b>5</b>  | 80<br>4 | 80<br>4  | 80<br>4                |         | 80<br>4                 | 0<br>0  | 80<br>4  | 20<br>1 | 40<br>2                | 40<br>2 | 40<br>2 | <b>100</b><br><b>5</b> | 80<br>4 |
|                             | K Self-monitor (behaviour)          | 13    | 54<br>7  | 38<br>5 | 46<br>6 | 85<br>11                | 54<br>7 | 69<br>9                 | 31<br>4 | 77<br>10 | 69<br>9                | 31<br>4 |                         | 31<br>4 | 69<br>9  | 23<br>3 | 38<br>5                | 23<br>3 | 31<br>4 | 62<br>8                | 62<br>8 |
|                             | L Self-monitor (outcome)            | 5     | 20<br>1  | 40<br>2 | 60<br>3 | 60<br>3                 | 40<br>2 | 60<br>3                 | 0<br>0  | 20<br>1  | 80<br>4                | 0<br>0  | 80<br>4                 |         | 60<br>3  | 20<br>1 | 60<br>3                | 60<br>3 | 40<br>2 | 20<br>1                |         |
|                             | M Feedback                          | 11    | 55<br>6  | 55<br>6 | 55<br>6 | 91<br>10                | 45<br>5 | 64<br>7                 | 36<br>4 | 55<br>6  | 73<br>8                | 36<br>4 | 82<br>9                 | 27<br>3 |          | 27<br>3 | 45<br>5                | 36<br>4 | 36<br>4 | 55<br>6                | 45<br>5 |
|                             | N Info where/when                   | 6     | 50<br>3  | 67<br>4 | 33<br>2 | 67<br>4                 | 67<br>4 | 67<br>4                 | 17<br>1 | 50<br>3  | 50<br>3                | 17<br>1 | 50<br>3                 | 17<br>1 | 50<br>3  |         | 50<br>3                | 33<br>2 | 33<br>2 | 33<br>2                | 50<br>3 |
|                             | O Instruct on how perform behaviour | 7     | 57<br>4  | 43<br>3 | 43<br>3 | 86<br>6                 | 43<br>3 | 57<br>4                 | 29<br>2 | 43<br>3  | 71<br>5                | 29<br>2 | 71<br>5                 | 43<br>3 | 71<br>5  | 43<br>3 |                        | 71<br>5 | 57<br>4 | 43<br>3                | 57<br>4 |
|                             | P Model behaviour                   | 5     | 60<br>3  | 40<br>2 | 60<br>3 | 80<br>4                 | 40<br>2 | 60<br>3                 | 40<br>2 | 20<br>1  | <b>100</b><br><b>5</b> | 40<br>2 | 60<br>3                 | 60<br>3 | 80<br>4  | 40<br>2 | <b>100</b><br><b>5</b> |         | 80<br>4 | 40<br>2                | 40<br>2 |
|                             | Q Prompt practice                   | 5     | 60<br>3  | 20<br>1 | 40<br>2 | 80<br>4                 | 20<br>1 | 60<br>3                 | 20<br>1 | 40<br>2  | <b>100</b><br><b>5</b> | 40<br>2 | 80<br>4                 | 60<br>3 | 80<br>4  | 40<br>2 | 80<br>4                | 80<br>4 |         | 40<br>2                | 60<br>3 |
|                             | R Follow-up prompts                 | 11    | 64<br>7  | 45<br>5 | 45<br>5 | 82<br>9                 | 55<br>6 | <b>100</b><br><b>11</b> | 36<br>4 | 64<br>7  | 55<br>6                | 45<br>5 | 73<br>8                 | 18<br>2 | 55<br>6  | 27<br>3 | 27<br>3                | 18<br>2 | 18<br>2 |                        | 55<br>6 |
|                             | S Social support                    | 9     | 56<br>5  | 33<br>3 | 44<br>4 | 89<br>8                 | 56<br>5 | 67<br>6                 | 33<br>3 | 78<br>7  | 56<br>5                | 33<br>4 | 89<br>8                 | 11<br>1 | 56<br>5  | 44<br>4 | 44<br>4                | 22<br>2 | 33<br>3 | 67<br>6                |         |

*Note.* The top number in the cells is the number of times that one feature co-occurred with the other as a percentage of the total number of times the feature was used. The lower number in italics is the total number of trials. Data in **bold** show when the pair of features co-occurred 100% of the time.

Table 2. Patterns and associations between intervention features used and effectiveness (standardised mean differences: SMDs)

| Trial                                   | SMD   | n    | Bias | Modes of Delivery       |                         |                         |                         | Behaviour Change Techniques |                         |                         |                         |                         |                             |                         |                            |                         |                         |                         |                         |                         |                         |                         |
|---|-------|------|------|-------------------------|-------------------------|-------------------------|-------------------------|-----------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-----------------------------|-------------------------|----------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
|   |       |      |      | Health<br>-care         | Phone                   | Home                    | Print                   | Model<br>beh                | Prompt<br>practice      | Feed-<br>back           | Graded<br>tasks         | Instruct<br>on beh      | Self-<br>monitor<br>outcome | Beh<br>goal<br>review   | Goal<br>setting<br>outcome | Self-<br>monitor<br>beh | Goal<br>setting<br>beh  | Follow-<br>up<br>prompt | Barrier<br>ID           | Social<br>support       | Info<br>where<br>when   | Info<br>conseq<br>indiv |
| Hertogh2010                             | 1.19  | 181  | high | X                       | X                       | X                       |                         | X                           | X                       | X                       | X                       | X                       |                             | X                       |                            |                         | X                       |                         |                         |                         |                         |                         |
| Babazono2007                            | 0.94  | 87   | low  | X                       |                         | X                       |                         |                             |                         | X                       | X                       |                         | X                           | X                       | X                          | X                       | X                       | X                       |                         |                         |                         |                         |
| Racette2008                             | 0.86  | 28   | high |                         |                         |                         |                         | X                           | X                       |                         | X                       | X                       |                             |                         |                            | X                       |                         |                         |                         | X                       |                         |                         |
| Kuller2006                              | 0.67  | 318  | high |                         |                         |                         |                         |                             |                         |                         |                         |                         |                             |                         |                            | X                       |                         |                         |                         |                         |                         |                         |
| Yates2009                               | 0.62  | 57   | low  | X                       |                         |                         |                         |                             | X                       | X                       |                         |                         |                             |                         | X                          | X                       | X                       | X                       |                         |                         |                         |                         |
| King2007                                | 0.39  | 123  | low  |                         | X                       | X                       |                         |                             | X                       | X                       |                         |                         | X                           |                         | X                          | X                       | X                       | X                       | X                       |                         |                         |                         |
| Lindstrom2003                           | 0.39  | 434  | low  | X                       | X                       | X                       | X                       | X                           | X                       | X                       |                         | X                       | X                           | X                       | X                          | X                       | X                       |                         |                         |                         |                         | X                       |
| McTiernan2007                           | 0.33  | 202  | low  | X                       |                         | X                       |                         | X                           | X                       | X                       |                         | X                       |                             |                         | X                          |                         |                         |                         | X                       |                         |                         |                         |
| Petrella2010                            | 0.25  | 329  | high | X                       | X                       | X                       |                         |                             |                         |                         |                         |                         |                             |                         |                            |                         |                         |                         |                         |                         |                         | X                       |
| Elley2003                               | 0.20  | 878  | low  | X                       | X                       | X                       | X                       |                             | X                       |                         |                         |                         |                             |                         |                            |                         | X                       | X                       |                         |                         | X                       |                         |
| Lawton2008                              | 0.19  | 1008 | low  | X                       | X                       | X                       |                         |                             |                         |                         |                         |                         | X                           | X                       | X                          | X                       | X                       | X                       | X                       |                         |                         | X                       |
| van Stralen2010                         | 0.19  | 920  | high |                         |                         | X                       | X                       |                             | X                       |                         |                         |                         |                             |                         | X                          | X                       | X                       | X                       | X                       | X                       | X                       | X                       |
| Opdenacker2008                          | 0.19  | 118  | low  |                         | X                       | X                       |                         |                             |                         |                         | X                       |                         |                             |                         |                            | X                       | X                       | X                       |                         |                         |                         |                         |
| van Keulen2011                          | 0.18  | 616  | high |                         | X                       | X                       | X                       |                             | X                       | X                       |                         |                         |                             |                         | X                          | X                       | X                       | X                       |                         |                         |                         | X                       |
| Martinson2010                           | 0.17  | 965  | low  |                         | X                       | X                       | X                       | X                           | X                       | X                       |                         | X                       | X                           | X                       | X                          | X                       | X                       | X                       | X                       | X                       | X                       | X                       |
| Harting2006                             | 0.10  | 1046 | high |                         | X                       | X                       |                         |                             |                         | X                       |                         |                         |                             |                         | X                          | X                       | X                       | X                       | X                       | X                       | X                       | X                       |
| Stoddard2004                            | 0.06  | 1075 | low  | X                       |                         |                         | X                       |                             |                         |                         |                         |                         |                             |                         |                            |                         |                         |                         | X                       |                         | X                       | X                       |
| Halbert2000                             | 0     | 299  | low  | X                       |                         |                         |                         |                             |                         | X                       |                         | X                       |                             | X                       | X                          | X                       | X                       | X                       |                         |                         | X                       | X                       |
| Werkman2010                             | -0.06 | 300  | low  |                         |                         | X                       | X                       |                             | X                       |                         | X                       |                         |                             | X                       | X                          | X                       |                         |                         |                         |                         |                         | X                       |
| Number of trials using feature (or not) |       |      |      | 10 (9)                  | 10 (9)                  | 14 (5)                  | 8 (11)                  | 5 (14)                      | 5 (14)                  | 11 (8)                  | 11 (8)                  | 7 (12)                  | 5 (14)                      | 5 (14)                  | 5 (14)                     | 13 (6)                  | 13 (6)                  | 11 (8)                  | 10 (8)                  | 9 (10)                  | 6 (13)                  | 10 (8)                  |
| SMD (95%CI) using feature               |       |      |      | 0.36<br>(0.19-<br>0.54) | 0.29<br>(0.16-<br>0.42) | 0.28<br>(0.17-<br>0.39) | 0.14<br>(0.07-<br>0.22) | 0.55<br>(0.18-<br>0.91)     | 0.55<br>(0.16-<br>0.95) | 0.40<br>(0.24-<br>0.56) | 0.40<br>(0.22-<br>0.59) | 0.42<br>(0.19-<br>0.65) | 0.43<br>(-0.04-<br>0.89)    | 0.33<br>(0.16-<br>0.50) | 0.28<br>(0.13-<br>0.42)    | 0.28<br>(0.15-<br>0.42) | 0.27<br>(0.11-<br>0.55) | 0.25<br>(0.13-<br>0.37) | 0.20<br>(0.11-<br>0.29) | 0.20<br>(0.11-<br>0.29) | 0.15<br>(0.08-<br>0.22) | 0.15<br>(0.09-<br>0.21) |
| SMD (95%CI) not using feature           |       |      |      | 0.23<br>(0.10-<br>0.36) | 0.32<br>(0.12-<br>0.51) | 0.37<br>(0.03-<br>0.70) | 0.48<br>(0.28-<br>0.67) | 0.22<br>(0.12-<br>0.32)     | 0.23<br>(0.13-<br>0.32) | 0.19<br>(0.05-<br>0.32) | 0.20<br>(0.09-<br>0.32) | 0.23<br>(0.12-<br>0.35) | 0.26<br>(0.17-<br>0.35)     | 0.28<br>(0.15-<br>0.42) | 0.30<br>(0.16-<br>0.44)    | 0.33<br>(0.13-<br>0.52) | 0.33<br>(0.15-<br>0.39) | 0.38<br>(0.18-<br>0.59) | 0.38<br>(0.17-<br>0.60) | 0.38<br>(0.18-<br>0.58) | 0.38<br>(0.21-<br>0.54) | 0.57<br>(0.32-<br>0.82) |
| Q test of subgroup differences P        |       |      |      | 0.213                   | 0.814                   | 0.647                   | 0.002                   | 0.094                       | 0.115                   | 0.042                   | 0.070                   | 0.160                   | 0.486                       | 0.673                   | 0.812                      | 0.731                   | 0.642                   | 0.270                   | 0.125                   | 0.118                   | 0.014                   | 0.001                   |

*Note.* Trials are ordered from most to least effective using the SMD. Modes of delivery and behaviour change techniques are grouped separately.

Within each group, features are ordered from left to right by the size of the difference in effect size estimate (SMD) between interventions using and not using each feature.

Healthcare = healthcare setting; Phone = telephone delivery; Home = home setting; Print = printed material delivery; Model beh = Model/

demonstrate the behaviour; Instruct on beh = Provide instruction on how to perform the behaviour; Barrier ID = Barrier identification/problem

solving; Social support = Plan social support/social change; Info where/when = Provide information on where and when to perform the behaviour; Info conseq indiv = Provide information on consequences of behaviour to the individual

IN PRESS

**Figure caption**

Figure 1. Forest plot of standardised mean differences (SMD) and 95% confidence intervals of the effect of interventions on physical activity behaviour compared with control conditions.

Studies are listed alphabetically by author.

IN PRESS